lead to techniques that do not necessarily facilitate safe operation and might even conflict with it. For example, in some application domains it is quite appropriate to permit real-time deadlines to be missed in a controlled way if that yields a useful improvement in resource utilization. In a safety-critical system this is unlikely to be acceptable.

With the needs of both technical areas in mind, the safety and real-time tracks concluded that much important research remains to be done:

- *Integrated scheduling analysis*

  In safety-critical systems, it is usually the case that certain deadlines are more significant than others. The hazard analysis performed for a given system will reveal precisely what actions need to be taken and when, and the results of hazard analysis include information about critical timing constraints. In dealing with issues such as real-time scheduling, the specific needs of the safety requirements of a system need to be taken into account.

- *Improved predictability*

  In considering the role of real-time analysis, the real-time and safety tracks agreed that new techniques are required to permit predictability of real-time performance in all system states. In this way, at least hazards will not arise as a result of inadequate real-time performance.

- *Integrated requirements analysis*

  Many real-time analyses are based on goals of maximizing factors such as utilization assuming average-case loads. In safety-critical systems, this is often not appropriate since it is most likely that worst-case analysis has to be assumed. This suggests that a research program in real-time systems that considers worst-case analysis systematically in support of safety-critical requirements would be very beneficial.

Since safety and real-time techniques are so closely related, it is clear that an integrated approach to analysis is needed. Research leading to an integrated framework of methods, tools, and techniques that addresses the interrelationship of the two areas is required.

systems operate with data that has to be protected. This data might range in sensitivity from relatively low such as patient record and treatment data in a medical database to relatively high such as targeting information in a weapons database. Clearly, the protection of sensitive information is a security issue.

A number of differences exist between the traditional domains of interest to the security research community and the domains that arise from consideration of safety-critical systems. The users of many safety-critical systems are specialists in other areas and not trained to operate or respect systems with security concerns. In a medical information system, for example, the users are unlikely to tolerate even simple security techniques such as physical barriers and access-control passwords.

A second difference is the need to deal with crisis situations in which the best interest of the community is served by controlled violation of the security mechanisms. Again using the medical information system as an example, it is likely that those requiring access to the information during a medical emergency will wish to bypass the security measures in the interest of speed. It is also possible that those needing the information are not routinely granted access yet the specific situation requires use of the information.

Several specific recommendations for research arose in the various discussions between the tracks:

- *Complex domains*

  A clear need was identified to develop techniques to support achievement of safety and security in complex application domains. Examples of where this technology is needed abound and include areas such as military weapons systems and many medical applications.

- *Unified modeling and analysis techniques*

  A second major research need is to develop unified approaches to modelling and analysis of safety and security properties in real systems. Both technical areas have modelling techniques that serve the specific needs of the area. But new issues arise when both qualities have to be present. It is easy to see, for example, that the implementation of security could affect safety adversely.

As well as noting the various issues raised by the overlap between security and safety, the joint discussions between the security and safety tracks revealed several opportunities for exploitation of techniques developed in one field by the other. This research could be fruitful quickly. An example is the application to safety systems of the kernel-based architecture that has been exploited successfully in security systems. A second example is the adaptation of hazard analysis from safety engineering to security systems with a view to exploiting the technology to improve threat analysis.

2.4 Safety and Real Time

Most safety-critical systems have real-time requirements. Thus it is often the case that temporal predictability and analysis are crucial to assuring safe operation. Inevitably, however, the goals of real-time systems when considered separately from safety tend to tional facility during normal operation. However, they permit the detection of certain types of faults and, in some cases after a fault has manifested itself, either allow the extent of the damage from the fault to be assessed or allow the effects of the fault to be masked.

The introduction of redundant elements into a system inevitably adds complexity and this leads to the very real possibility of degraded safety performance. The Sizewell B primary protection system, for example, includes more than 600 microprocessors most of which are present to provide redundant operation. It is difficult to see how such a system can be safer than one in which the total number of processors was reduced even if this meant some faults would be detected rather than masked.

A similar example was raised during the discussions between the safety and fault-tolerance tracks in which concerns were expressed about the complexity of modern asynchronous, multi-channel, avionics architectures. It is not clear that systems of this complexity can be analyzed adequately with existing techniques.

With these and other examples in mind, the safety and fault-tolerance tracks concluded that research into modelling and analysis techniques needs to be undertaken. Specific analyses for which new techniques are required include:

- *Analysis of asynchronous systems.*

  Analysis techniques that predict the performance of complex architectures, especially asynchronous architectures, is required. This issue arises because of the redundancy that has to be present in fault-tolerant systems. Confidence in such architectures is essential if the safety of the associated systems is to be demonstrated.

- *Trade-off analysis.*

  The ability to make trade-offs between the reliability, availability, and safety requirements in complex digital systems is essential. If reliability requirements dictate levels of redundancy that raise safety concerns then a trade-off needs to be made. Similarly, if safety requirements dictate a system architecture that will lead to unacceptable availability then the architecture needs to be reconsidered.

- *Implication of interaction.*

  Different techniques are required to deal with different types of faults in fault-tolerant systems. It is likely that these techniques interact in ways that are not obvious and that could easily lead to hazardous system states. Analysis techniques that permit the determination of the implications of interacting fault tolerance techniques are required.

2.3 Safety and Security

Security is a necessary part of maintaining safety in certain systems. All the care in the world in system development and operation can be undone by an attack undertaken by a user with malicious intent. Such an attack is most likely to be launched by an unauthorized user and the prevention of unauthorized access is a security issue.

As well as malicious attacks that lead to hazards, it is often the case that safety-critical that research be conducted into the interface between systems and software engineering with a view to achieving an integration and merging of systems-level techniques into and with software engineering.

- *Improved Software Architectures*

  To promote the achievement of safe operation, much can be done by using appropriate software structures. For example, in a safety-critical system if a number of conditions have to be checked and all found to be true in order for some possible hazardous action to take place, it is preferable that the initial values of the conditions maintained by the software be false. In this way, if any disruption in processing occurs for whatever reason, there is no danger of the action taking place if the necessary conditions have not been checked. This type of approach needs to be developed further to produce software architectures that are based on safety goals.

- *Human/Computer Interface*

  It is clear from accident investigations that the cause of many mishaps involving computer-based systems lies in the human-computer interface. Interfaces are frequently confusing, overly complex, difficult to use, and so on. Systematic research into the design of user interfaces that promote safe operation is a high priority.

- *Hardware/Software Interaction*

  It is often the case that a simple modification or addition to the hardware in a safety-critical system can avoid a complex (and probably erroneous) software solution to a safety concern. A well-known example is the Therac 25 therapy system in which a simple hardware interlock could have prevented the fatal radiation overdoses that were delivered. A system-level approach to safety that promotes hardware/software synergy in design can frequently provide elegant solutions. The applicability of such synergy though obviously valuable has not been the subject of systematic study. It should be the subject of a thorough research program.

- *Development and Management Processes*

  Many accidents are attributable to the use of poor development and/or management processes. Although the technology exists in some cases to support the effective development of safety-critical systems, these technologies are frequently not employed or are applied improperly. The development of rigorous processes that can be applied dependably is an area of research that should be pursued.

2.2 Safety and Fault Tolerance

In almost any safety-critical system, faults that arise such as the degradation of a physical device must be addressed so as to permit a safe level of service to be maintained. This level of service might be limited to an orderly shutdown of the equipment in what are called “fail-safe” systems. From the perspective of safety, therefore, fault tolerance is a technology that permits crucial safety issues to be addressed.

Much of the available technology in the area of fault tolerance depends for its performance on redundancy. Supplements are added to systems that do not provide any addi-

- *Realistic Applications*

  We recommend that the funding agencies solicit proposals for and provide funding for research efforts applying advanced high-assurance techniques to large-scale applications. Such large scale efforts would include shadow efforts of ongoing industrial developments or realistic problems provided by industry or government agencies (e.g., nuclear safety shutdown systems). Requests for proposals should solicit research in applying advanced safety analysis and software engineering techniques to such systems to address multiple properties such as safety and real-time constraints.

- *Industrial-strength Tools And Techniques*

  We recommend that funding be directed to support (prototype) development of improved tools supporting engineering and analysis of systems with multiple high-assurance properties. To ensure such tools scale to industrial problems, they must be developed for and tested on realistic applications.

- *Empirical Evaluation*

  We recommend that funding be provided for empirical (experimental) validation of new technology. There is currently a lack of objective evidence for the relative effectiveness of new techniques or tools - this leaves industrial program managers little basis to support acquisition of advanced software technology. To help validate the models used and to help refine the research issues to be addressed, careful data collection from the field should be undertaken over extended periods and made available to the community for analysis.

- *Long-term Perspective*

  We recommend that funding agencies take long-term perspective on their research programs. To allow adequate time for communication and assimilation of issues and technology between research and industry, the funding agencies need, where possible, to support long term (2-5+ years) research and technical transfer efforts. Long-term funding cycles need to be established for researchers and, where there is industrial involvement, consideration needs to be given to providing funding to the industrial participants.

- *Unified Approach*

  We recommend that funding agencies increase support for research efforts that addres multiple high-assurance properties. Efforts should include soliciting proposals to develop modeling and analysis techniques that proved a unified approach to more than one of the safety, security, fault-tolerance, and real-time properties.

2.1 Safety

In the area of safety considered alone, the safety track concluded that research is required in a number of areas. It is presently the case that systems engineers and software engineers tend to operate with different techniques yet they must communicate in order to ensure that the necessary analysis is performed on the software. It is essential, therefore

This is essential not only to ensure continued operation in the very short term but also to permit the necessary repair to be scheduled.

Finally, the majority of the complexity in this system lies in the external “programmer”. This is a device that is used to communicate with the implanted unit, and through which operating parameters are set and therapy history is acquired and displayed. The programmer might be connected to a network and certainly contains sensitive patient data that must be protected. Operation of the programmer by unqualified personnel poses a potential hazard to the patient so access control must be effected to prevent unauthorized use. All of these topics amount to a requirement for *security* in the system.

Perhaps surprisingly, this application has requirements in all four areas that the workshop addressed. It is just one of many examples that were discussed that also exhibited this property of requiring multiple qualities.

2 Research Recommendations

A recurring theme in both the discussion held by the safety track alone and with members of the other tracks was the need for improved technology development and transfer focused on the problems of developing large-scale high-assurance systems. There is strong evidence that the degree of communication between the research and the application communities is not as effective as it needs to be. The application community tends to raise issues that need to be resolved and the research community generates new results that are expected to produce benefits but the following difficulties remain:

- There are a large number of open issues in industry that are not being addressed by the research community.
- Research results and technology produced by the research community are often not being applied by the application community.
- It is often not clear to practitioners how a particular new technology should be applied.
- In many cases, practitioners are unaware of new technologies.
- Technology that is developed by researchers is frequently not developed to industrial strength.
- The funding and reward structures in the research community do not promote industrial-strength technology development and transfer.

The safety track’s primary conclusion was that new research funding should be directed to close the gap between research and practice, addressing, in particular, the issues above. The recommended areas of funding are:

- *Industry/Academic/Government Partnerships*

  We recommend that funding be directed to research efforts that partner industry, Government, and academia. In particular, efforts where industry defines the problem, research develops potential solutions accounting for industrial constraints, and the results are demonstrated as a cooperative effort.

The detection module immediately notifies the Bradycardia therapy module when a slow heart beat is detected. The detection module counts a physician programmed number of fast beats before notifying the Tachycardia therapy module, this prevents a single or short number of fast beats from being recognized as a life threatening episode.

- *Bradycardia Therapy*

  The Bradycardia therapy module is responsible for delivering a low energy pacing pulse to the patient's heart when a slow beat is detected. This provides the patient with a minimum life sustaining heart rate. The physician programs the peak voltage of the pulse (amplitude) and the duration of the pulse (pulse width).

- *Tachycardia Therapy*

  The Tachycardia therapy module is responsible for delivering a high energy shock to the patient's heart when a fast arrhythmia episode is detected. If the rate was classified as Tachy then a physician-programmable shock is delivered to the patient. If the rate was classified as Fib, then a maximum energy (35 joule) shock is delivered to the patient. The physician may program the leading edge polarity (+ or -), the wave form (mono-phasic or bi-phasic), and the Tachy shock energy (0.1 to 35 joules).

- *Therapy History*

  The therapy history module is responsible for maintaining a rotating buffer of past detection and therapy attempts. The buffer contains a maximum of the last 16 therapy attempts.

Clearly, such a system must be considered safety critical. Some of the hazards that can arise with this system are:

- Incorrect diagnosis reached based on stored history data.
- Arrhythmia accelerated by delivery of therapy.
- Therapy delivered but ineffective or delivered inappropriately.
- Therapy not enabled, specifically inhibited, or not delivered for some reason.
- Fractured or dislodged sensing lead.

Required System Qualities

An implanted defibrillator is an example of a system requiring more than one of the qualities being addressed by the workshop. Clearly, there is an overriding concern for *safety*. It is essential that the device not operate in a fashion that leads to a hazard for the patient.

But, given the application, it is also clear that the system has stringent *real-time* characteristics. Much of the analysis that has to be performed during diagnosis is time-dependent as are any therapeutic actions that have to be taken.

An implanted defibrillator is very difficult to service, and so it must be *fault tolerant*.

range by monitoring the reaction products in real time (reactant supply has to be adjusted to deal with changes in the state of the catalyst, build up of reaction by products, etc.),

- detection of failure that could lead to an undesirable change in the reaction conditions by monitoring of all peripheral equipment (pumps, valves, other actuators) in real time,
- coping with hardware failures of the computing platform itself.

As is seen regularly, the above list of requirements can be difficult to meet. Chemical reactor accidents are common but rarely reported except in circumstances such as Bohpal or a major refinery accident.

The issue with chemical reactors is almost entirely one of *safety*. A reactor that is idle is safe but not earning revenue. This is perfectly safe. But a reactor that fails to adjust the relative amount of reactants in a reaction vessel can quickly enter an unstable operating regime and release far more energy than the reaction vessel can contain. The results is usually a major spill of hot, toxic materials and the destruction of expensive equipment.

In designing and building a chemical reactor facility, the activities undertaken are the classic techniques of safety engineering. Yet if the facility is to be controlled with a digital system, the question remains as to how to engineer the software so as to contribute positively to plant safety. Interestingly, the chemical reactor industry is not regulated by any Government agency with the same degree of thoroughness that is effected by say the NRC, the FDA or the FAA.

Example System with Multiple Properties - Heart Defibrillator

In the USA about 400,000 people die each year from Sudden Cardiac Death Syndrome. Implantable defibrillators are used to treat people with Sudden Cardiac Death Syndrome. These patients have generally died if they were not fitted with this device and are very dependent on the device for keeping them live.

The defibrillator is required to take therapeutic input from a physician and provide therapy history to the physician. The device must sense and treat Bradycardia (slow heart rhythms) and Tachycardia (fast heart rhythms) and store information about how the arrhythmias were detected and treated. The device can be functionally decomposed into the following five modules:

- *Communications*

  The communication module is responsible for reliably transferring the physician's programmable parameters to the detection and therapy modules and transferring the history module's data to the physician.

- *Detection*

  The detection module classifies each beat according to the physician programmed heart rate limits. A beat is classified as Bradycardic if it is below the Brady rate limit, as Tachycardic if it is above the Tachy limit and as Fibrillation if it is above the Fib limit.

Definitions

The following are generally accepted definitions of the terms used in discussing safety:

Definition: **Accident**
An accident is an undesirable event which results in unacceptable consequences.

Definition: **Hazard**
A hazard is a system state that if left uncorrected could lead to an accident.

Definition: **Risk**
The risk associated with a hazard is a product of the cost associated with an accident and its probability of occurrence.

Definition: **Reliability**
The conditional probability R(t) that a system will operate without failure in a specified environment for the interval [0, t] given that it was operational at time 0.

Definition: **Availability**
The probability A(t) that a system is operational at time t.

Definition: **Safety**
Assurance that the system will operate within a specified environment without resulting in unacceptable risk.

Some surprising combinations of properties can exist. It is possible for a system to be highly available yet unreliable if the system fails frequently but restarts itself with minimal delay. Similarly, a system does not have to be reliable nor available to be safe. In fact, in the limit, a system that is not operational is neither reliable nor available but is probably safe.

By contrast, it is quite possible to have a system that is reliable yet unsafe. This last situation might occur, for example, in a system that meets its requirements but causes accidents because the requirements were not appropriate.

Safety is concerned with systems that must avoid accidents. Such safety-critical applications include systems that can cause accidents (e.g., an autopilot) and systems intended to prevent accidents (e.g., reactor safety-shutdown system).

As an example, consider a chemical reactor system that is producing a product using a reaction that operates at high temperature, that is itself a source of energy, and whose reactants are toxic. Adequate management of such a system requires as a minimum:

- operating the energy control facilities of the reactor correctly by monitoring the reactant temperature in real time and adjusting the necessary actuators (sometimes the reaction has to be heated and sometimes cooled),
- adjustments of the supply of reactants to keep the reaction operating in the optimal

text, and demonstrate these solutions on realistic problems.

- *Industry/research cooperation* - provide support for cooperative research efforts between industry and research centers; in particular, proposals that provide for effective transition to and application by industry.

The body of the report gives background on safety issues and provides detailed research recommendations especially as related to issues in developing safety-critical systems.

1.2 Approach

The safety track discussed the major issues facing the safety field and then considered the interaction between safety and the other tracks. We concluded that the four areas which were under consideration (real time, fault tolerance, security, and safety) interrelate in two entirely different ways:

- *Overlap*

  There are systems that have requirements from more than one area. The discussion focused on how these requirements can be met.

- *Exploitation*

  Techniques have been developed within each of the areas that help deal with problems specific to that area. The discussion focused on how these techniques be exploited in different areas.

Detailed recommendations were developed addressing how best to exploit each type of relationship.

1.3 Background

Safety is a property that systems have that can be summarized informally as “the system does not do any unintentional harm.” Examples of harm that a system might do include endangerment of human life or the loss of valuable equipment or data. Systems where safety is important are often referred to as *safety-critical*.

Examples of safety-critical systems abound. Obviously, many defense systems are safety-critical. The harm that can be done by the inappropriate deployment of a weapon is immense. Many transportation systems are safety-critical also. Clearly, the flight control system in a commercial air transport is safety-critical as is the air-traffic-control system.

Surprisingly, the range of safety-critical systems that affect modern society is much wider than one might think. Many computerized financial systems can be thought of as safety-critical because the harm that could be done to the economy would be immense if such systems were to fail. Other areas of importance are the electronic telephone and other telecommunications networks. Societal dependence on all aspects of the telephone system is significant. Loss of local service, for example, immediately removes the capability to summon emergency services. Similarly, loss of long-distance service can cause serious disruption to business activities and an ensuing loss of revenue.

1 INTRODUCTION

1.1 Summary

The safety track determined that new funding directions are needed to ensure that the research community will meet the future needs of industry for reliable and cost-effective development of high-assurance systems. The safety track participants worked with participants in the fault-tolerance, security, and real-time tracks to identify areas where current research or research funding does not adequately address industry and government needs for improved software technology. Discussions included representatives from Government funding organizations, research institutions, and industrial developers of high-assurance systems.

The safety track found that there is currently a significant gap between the technologies needed by industry to reliably develop high-assurance systems and the technical results being provided by the research community. In many cases, research is not being exploited by developers and significant technical problems are not be addressed by researchers. In particular, this is true of the increasing numbers of systems where technology in two or more of the high-assurance areas of safety, fault-tolerance, security, or real-time are needed. In summary, we identified the following shortcomings:

- *Lack of integrated technologies* - Current research typically focuses on problems in one of the areas of safety, fault-tolerance, security, or real-time. Increasingly, systems are being developed with multiple high-assurance properties (e.g., medical systems with safety, real-time, and fault-tolerance requirements). Current software engineering technology does provide adequate development, modeling, and analysis techniques for such systems.
- *Lack of industrial relevance* - There is a lack of communication between industrial developers and the research community. Researchers do not adequately understand or address industry's problems and constraints. As a result, industry does not perceive much of the technology as addressing their needs (e.g., in ability to scale or cost effectiveness).
- *Lack of transition support* - Technology transition remains a significant obstacle. Support is lacking for demonstrating the industrial relevance of research results, adapting new technology to industrial use, or inserting such technology into industrial practice.

The safety track identified several research and technology transfer areas where carefully directed funding can improve public safety, increase U.S. industry competitiveness, and reduce government costs for safety-critical systems. We recommend that the funding agencies solicit proposals for and provide support in the following areas:

- *Unified approaches* - provide support for work developing unified models and methods in two or more of the areas of safety, fault-tolerance, security, and real-time.
- *Industrially-relevant research* - provide support for research efforts that address real industrial problems, seek scalable solutions applicable in the industrial con-

WORKSHOP
ON
HIGH ASSURANCE COMPUTING

SAFETY TRACK REPORT

John C. Knight (Chair), University of Virginia
Mike Blackwell, Carnegie Mellon University
Kevin Driscoll, Honeywell
Lynn Elliott, Cardiac Pacemakers Inc.
Stuart Faulk, NRL
Nancy Leveson, University of Washington
Gary Preckshot, Lawrence Livermore Lab.
Roger Schultz, Rockwell Collins Commercial Avionics
Dolores Wallace, NIST
Chuck Weinstock, SEI
Janusz Zalewski, Embry-Riddle Aeronautical Univ.